02_model_training

October 25, 2025

[1]:
```
!pip install GPUtil contractions
```

```
Collecting GPUtil
  Downloading GPUtil-1.4.0.tar.gz (5.5 kB)
  Preparing metadata (setup.py) … done
Collecting contractions
  Downloading contractions-0.1.73-py2.py3-none-any.whl.metadata (1.2 kB)
Collecting textsearch>=0.0.21 (from contractions)
  Downloading textsearch-0.0.24-py2.py3-none-any.whl.metadata (1.2 kB)
Collecting anyascii (from textsearch>=0.0.21->contractions)
  Downloading anyascii-0.3.3-py3-none-any.whl.metadata (1.6 kB)
Collecting pyahocorasick (from textsearch>=0.0.21->contractions)
  Downloading pyahocorasick-2.2.0-cp311-cp311-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (13 kB)
Downloading contractions-0.1.73-py2.py3-none-any.whl (8.7 kB)
Downloading textsearch-0.0.24-py2.py3-none-any.whl (7.6 kB)
Downloading anyascii-0.3.3-py3-none-any.whl (345 kB)
                         345.1/345.1 kB
9.5 MB/s eta 0:00:00
Downloading
pyahocorasick-2.2.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl
(113 kB)
                         113.9/113.9 kB
8.9 MB/s eta 0:00:00
Building wheels for collected packages: GPUtil
  Building wheel for GPUtil (setup.py) … done
  Created wheel for GPUtil: filename=GPUtil-1.4.0-py3-none-any.whl size=7392
sha256=800850e8e31eb54c5035bdb49c5bbdec681f1592548dde6fe1ca7ca98ebbceba
  Stored in directory: /root/.cache/pip/wheels/2b/4d/8f/55fb4f7b9b591891e8d3f729
77c4ec6c7763b39c19f0861595
Successfully built GPUtil
Installing collected packages: GPUtil, pyahocorasick, anyascii, textsearch,
contractions
Successfully installed GPUtil-1.4.0 anyascii-0.3.3 contractions-0.1.73
pyahocorasick-2.2.0 textsearch-0.0.24
```

[2]:
```
# ===== Set up project ======
import os
```

```
import shutil

# Copy source code to working directory
src_path = r"/kaggle/input/machine-learning-resource/"
dst_path = r"/kaggle/working/machine-learning-resource/"

shutil.copytree(src_path, dst_path, dirs_exist_ok=True)
```

[2]:
```
'/kaggle/working/machine-learning-resource/'
```

[3]:
```
# Add root directory to PYTHON path
import sys
from pathlib import Path

# === LOCAL ===
# root_dir = str(Path.cwd().parent.parent.absolute())
# if not root_dir in sys.path:
#     sys.path.insert(0, root_dir)

# === KAGGLE ===
root_dir = "/kaggle/working/machine-learning-resource"
if not root_dir in sys.path:
    sys.path.insert(0, root_dir)
```

1 Model Training - LSTM & BiLSTM

[4]:
```
import pandas as pd
import matplotlib.pyplot as plt

from config import Config
from src.utils.gpu_utils import GPUMemoryManager
from src.data.preprocessing import DataPreprocessor
from src.models.bilstm_attention import BiLSTMAttentionModel
from src.models.lstm_attention import LSTMAttentionModel
from src.training.trainer import ModelTrainer
from src.utils.helpers import save_tokenizer
```

```
2025-10-23 17:52:17.105981: E
external/local_xla/xla/stream_executor/cuda/cuda_fft.cc:477] Unable to register
cuFFT factory: Attempting to register factory for plugin cuFFT when one has
already been registered
WARNING: All log messages before absl::InitializeLog() is called are written to
STDERR
E0000 00:00:1761241937.302717      19 cuda_dnn.cc:8310] Unable to register cuDNN
factory: Attempting to register factory for plugin cuDNN when one has already
been registered
E0000 00:00:1761241937.352359      19 cuda_blas.cc:1418] Unable to register
```

```
cuBLAS factory: Attempting to register factory for plugin cuBLAS when one has
already been registered
```

[5]:
```python
# Copy outputs from previous version to working directory
# (Used to continue training the model if the allowed session runs out)
input_dirs = [
    d for d in os.listdir('/kaggle/input/') if 'machine-translation-project' in␣
 ↪d
]

if input_dirs:
    input_dir = f"/kaggle/input/{input_dirs[0]}/machine-learning-resource/"
    work_dir = "/kaggle/working/machine-learning-resource/"

    os.makedirs(input_dir, exist_ok=True)
    shutil.copytree(
        input_dir,
        work_dir,
        ignore=shutil.ignore_patterns("src"),
        dirs_exist_ok=True
    )

    print("Loaded outputs from previous version")
else:
    print("No previous outputs found. Starting fresh.")
```

```
Loaded outputs from previous version
```

1.1 1. GPU Setup

[6]:
```python
GPUMemoryManager.clear_session()
GPUMemoryManager.setup_gpu(
    memory_limit_mb=Config.GPU_MEMORY_LIMIT,
    allow_growth=Config.GPU_MEMORY_GROWTH
)
if Config.USE_MIXED_PRECISION:
    GPUMemoryManager.enable_mixed_precision()
    GPUMemoryManager.get_memory_info()
```

```
Session cleared
GPU memory growth enabled
GPU memory limited to 15000 MB
Found 1 GPU(s)
Mixed precision enabled: mixed_float16
Compute dtype: float16
Variable dtype: float32
GPU 0: 3 MB / 16384 MB (0.0%)
```

1.2 2. Configuration

[7]:
```
config = Config.to_dict()

print("Configuration:")
for key, value in config.items():
    print(f"   {key}: {value}")

print("\nMemory Estimate:")
for key, value in Config.estimate_memory().items():
    print(f"   {key}: {value}")
```

```
Configuration:
   batch_size: 128
   epochs: 100
   validation_split: 0.1
   learning_rate: 0.001
   beta_1: 0.9
   beta_2: 0.98
   optimizer_epsilon: 1e-09
   optimizer_dynamic: True
   optimizer_initial_scale: 32768
   optimizer_dynamic_growth_steps: 2000
   use_lr_scheduler: True
   warmup_steps: 8000
   total_steps: 100000
   min_lr: 1e-07
   early_stopping_patience: 5
   reduce_lr_patience: 3
   reduce_lr_factor: 0.5
   save_best_only: True
   monitor: val_loss
   tensorboard_update_freq: epoch
   max_vocab_size_src: 30000
   max_vocab_size_trg: 25000
   min_word_frequency: 1
   embedding_dim: 64
   lstm_units: 128
   attention_heads: 4
   layer_dropout: 0.2
   lstm_dropout: 0.2
   attention_dropout: 0.1
   ln_epsilon: 1e-06
   max_length_src: 40
   max_length_trg: 50
   use_mixed_precision: True
   gpu_memory_growth: True
   gpu_memory_limit: 15000
```

```
   label_smoothing: 0.05
   use_layer_norm: True
   beam_width: 5
   use_beam_search: True

Memory Estimate:
   total_params: 10,444,288
   model_memory_mb: 59.76 MB
   batch_memory_mb: 15625.00 MB
   total_memory_gb: 15.32 GB
```

1.3 3. Data Preprocessing

[8]:
```python
# Tokenizer
tokenizer_path = f"{Config.ARTIFACT_PATH}/tokenizers"
processed_data_path = f"{Config.DATA_PATH}/processed/processed_df.csv"

preprocessor = DataPreprocessor(
    max_vocab_src=Config.MAX_VOCAB_SIZE_SRC,
    max_vocab_trg=Config.MAX_VOCAB_SIZE_TRG,
    min_frequency=Config.MIN_WORD_FREQUENCY,
    expand_contractions=True,
    remove_punctuation=True,
    name_logger='data_preprocessing',
    filename_logger=f'{Config.LOG_DIR}/data_preprocessing.log'
)

if not os.path.exists(processed_data_path):
    print("Processing data from scratch...")

    # Load data
    df = preprocessor.load_data(
        src_path=f"{Config.DATA_PATH}/raw/en.txt",
        trg_path=f"{Config.DATA_PATH}/raw/vi.txt",
        max_length_src=Config.MAX_LENGTH_SRC,
        max_length_trg=Config.MAX_LENGTH_TRG
    )

    # Save processed data
    os.makedirs(os.path.dirname(processed_data_path), exist_ok=True)
    df.to_csv(processed_data_path, index=False)
    print(f"Saved processed data to {processed_data_path}")
else:
    print("Loading cached processed data...")
    df = pd.read_csv(processed_data_path, encoding="utf-8")

print(f"Dataset: {df.shape}")
```

```
Loading cached processed data…
Dataset: (143144, 2)
```

[9]:
```python
# Split
train_df, val_df, test_df = preprocessor.split_data(df)
print(f"Train: {len(train_df)}, Val: {len(val_df)}, Test: {len(test_df)}")

tokenizer_en_path = f'{tokenizer_path}/tokenizer_en.pkl'
tokenizer_vi_path = f'{tokenizer_path}/tokenizer_vi.pkl'

if not os.path.exists(tokenizer_en_path):
    print("Building tokenizers from scratch...")
    tokenizer_en, tokenizer_vi = preprocessor.build_tokenizers(train_df)

    # Save
    os.makedirs(tokenizer_path, exist_ok=True)
    save_tokenizer(tokenizer_en, tokenizer_en_path)
    save_tokenizer(tokenizer_vi, tokenizer_vi_path)
    print(f"Saved tokenizers to {tokenizer_path}")
else:
    print("Loading cached tokenizers...")
    from src.utils import load_tokenizer
    tokenizer_en = load_tokenizer(tokenizer_en_path)
    tokenizer_vi = load_tokenizer(tokenizer_vi_path)
    preprocessor.tokenizer_src = tokenizer_en
    preprocessor.tokenizer_trg = tokenizer_vi

print(f"English vocab size: {len(tokenizer_en.word_index)}")
print(f"Vietnamese vocab size: {len(tokenizer_vi.word_index)}")
```

```
Train: 114515, Val: 14314, Test: 14315
Loading cached tokenizers…
Tokenizer loaded from …/artifacts/tokenizers/tokenizer_en.pkl
Tokenizer loaded from …/artifacts/tokenizers/tokenizer_vi.pkl
English vocab size: 36621
Vietnamese vocab size: 27045
```

[10]:
```python
# Prepare sequences
en_train, vi_in_train, vi_out_train = preprocessor.prepare_sequences(
    data=train_df,
    max_len_src=Config.MAX_LENGTH_SRC,
    max_len_trg=Config.MAX_LENGTH_TRG
)

en_val, vi_in_val, vi_out_val = preprocessor.prepare_sequences(
    data=val_df,
    max_len_src=Config.MAX_LENGTH_SRC,
    max_len_trg=Config.MAX_LENGTH_TRG
```

```
)

print(f"Training sequences: {en_train.shape}")
```

```
2025-10-23 17:53:01 - data_preprocessing - INFO - Sequences memory: 61.16 MB
2025-10-23 17:53:02 - data_preprocessing - INFO - Sequences memory: 7.64 MB
Training sequences: (114515, 40)
```

1.4 4. Build BiLSTM Model

[11]:
```
model_builder = BiLSTMAttentionModel(
    config=Config,
    name_logger="bilstm_attention",
    filename_logger=f"{Config.LOG_DIR}/bilstm_attention.log"
)
bilstm_model = model_builder.build(
    vocab_size_src=Config.MAX_VOCAB_SIZE_SRC,
    vocab_size_trg=Config.MAX_VOCAB_SIZE_TRG,
    max_len_src=Config.MAX_LENGTH_SRC,
    max_len_trg=Config.MAX_LENGTH_TRG
)

bilstm_model.summary()
```

```
I0000 00:00:1761241982.380087      19 gpu_device.cc:2022] Created device
/job:localhost/replica:0/task:0/device:GPU:0 with 15000 MB memory:  -> device:
0, name: Tesla P100-PCIE-16GB, pci bus id: 0000:00:04.0, compute capability: 6.0
```

Model: "bilstm_attention_model"

Layer (type)	Output Shape	Param #	Connected to
encoder_input (InputLayer)	(None, 40)	0	-
encoder_embedding (Embedding)	(None, 40, 64)	1,920,000	encoder_input[0]…
decoder_input (InputLayer)	(None, 50)	0	-
encoder_ln1 (LayerNormalizatio…	(None, 40, 64)	128	encoder_embeddin…
decoder_embedding (Embedding)	(None, 50, 64)	1,600,000	decoder_input[0]…

encoder_dropout (Dropout)	(None, 40, 64)	0	encoder_ln1[0][0]
not_equal (NotEqual)	(None, 40)	0	encoder_input[0]…
decoder_ln1 (LayerNormalizatio…	(None, 50, 64)	128	decoder_embeddin…
encoder_bilstm (Bidirectional)	[(None, 40, 256), (None, 128), (None, 128), (None, 128), (None, 128)]	197,632	encoder_dropout[… not_equal[0][0]
decoder_dropout (Dropout)	(None, 50, 64)	0	decoder_ln1[0][0]
concatenate (Concatenate)	(None, 256)	0	encoder_bilstm[0… encoder_bilstm[0…
concatenate_1 (Concatenate)	(None, 256)	0	encoder_bilstm[0… encoder_bilstm[0…
decoder_lstm (LSTM)	[(None, 50, 256), (None, 256), (None, 256)]	328,704	decoder_dropout[… concatenate[0][0… concatenate_1[0]…
decoder_ln2 (LayerNormalizatio…	(None, 50, 256)	512	decoder_lstm[0][…
encoder_ln2 (LayerNormalizatio…	(None, 40, 256)	512	encoder_bilstm[0…
not_equal_1 (NotEqual)	(None, 50)	0	decoder_input[0]…
attention_layer (MultiHeadAttentio…	(None, 50, 256)	526,080	encoder_ln2[0][0… not_equal[0][0], decoder_ln2[0][0… not_equal_1[0][0… encoder_ln2[0][0… not_equal[0][0]
attention_ln (LayerNormalizatio…	(None, 50, 256)	512	attention_layer[…
residual_add (Add)	(None, 50, 256)	0	decoder_ln2[0][0…

```
                                                                 attention_ln[0][…

 residual_ln           (None, 50, 256)            512   residual_add[0][…
 (LayerNormalizatio…

 concat_output         (None, 50, 512)              0   residual_ln[0][0…
 (Concatenate)                                          attention_ln[0][…

 concat_dropout        (None, 50, 512)              0   concat_output[0]…
 (Dropout)

 cast (Cast)           (None, 50, 512)              0   concat_dropout[0…

 decoder_dense         (None, 50, 25000)   12,825,000   cast[0][0]
 (Dense)


 Total params: 17,399,720 (66.37 MB)

 Trainable params: 17,399,720 (66.37 MB)

 Non-trainable params: 0 (0.00 B)
```

[12]:
```
for layer in bilstm_model.layers[:5]:
    print(f"{layer.name}: dtype={layer.dtype}, compute_dtype={layer.
 ↪compute_dtype}")
```

```
encoder_input: dtype=int32, compute_dtype=float16
encoder_embedding: dtype=float32, compute_dtype=float32
decoder_input: dtype=int32, compute_dtype=float16
encoder_ln1: dtype=float32, compute_dtype=float16
decoder_embedding: dtype=float32, compute_dtype=float32
```

1.5 5. Train BiLSTM

[13]:
```
steps_per_epoch = len(en_train) // Config.BATCH_SIZE
Config.TOTAL_STEPS = steps_per_epoch * Config.EPOCHS
model_name = "bilstm"

print(f"Total training steps: {Config.TOTAL_STEPS}")
print(f"Steps per epoch: {steps_per_epoch}")
print(f"Batch size (global): {Config.BATCH_SIZE}")

trainer = ModelTrainer(
    model=bilstm_model,
    config=Config,
```

```
    model_name=model_name,
    logger_name="bilstm_attention",
    logger_file=f"{Config.LOG_DIR}/bilstm_attention.log"
)

bilstm_history = trainer.train(
    train_data=(en_train, vi_in_train, vi_out_train),
    val_data=(en_val, vi_in_val, vi_out_val),
    resume=True
)
```

```
Total training steps: 89400
Steps per epoch: 894
Batch size (global): 128
2025-10-23 17:53:03 - bilstm_attention - INFO - Loading checkpoint:
/kaggle/working/machine-learning-
resource/artifacts/bilstm/checkpoints/checkpoint_epoch_070.keras
2025-10-23 17:53:05 - bilstm_attention - INFO - Resuming from epoch 70
2025-10-23 17:53:05 - bilstm_attention - INFO - Mixed precision optimizer
enabled with increased stability
2025-10-23 17:53:05 - bilstm_attention - INFO - Model compiled:
2025-10-23 17:53:05 - bilstm_attention - INFO - Loss: Label Smoothing ( =0.05)
2025-10-23 17:53:05 - bilstm_attention - INFO - LR Schedule: True
2025-10-23 17:53:05 - bilstm_attention - INFO - Training started:
2025-10-23 17:53:05 - bilstm_attention - INFO - Batch size: 128
2025-10-23 17:53:05 - bilstm_attention - INFO - Initial epoch: 70
2025-10-23 17:53:05 - bilstm_attention - INFO - Total epochs: 100
2025-10-23 17:53:05 - src.training.callbacks - INFO - Loaded existing history
with 77 epochs
Epoch 71/100
895/895             0s 589ms/step -
accuracy: 0.9266 - loss: 1.0458
Epoch 71: val_loss improved from inf to 1.13494, saving model to
/kaggle/working/machine-learning-
resource/artifacts/bilstm/checkpoints/best_model.keras
2025-10-23 18:02:44 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895             578s 617ms/step -
accuracy: 0.9266 - loss: 1.0458 - val_accuracy: 0.9161 - val_loss: 1.1349
Epoch 72/100
895/895             0s 598ms/step -
accuracy: 0.9270 - loss: 1.0456
Epoch 72: val_loss did not improve from 1.13494
2025-10-23 18:12:00 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895             556s 622ms/step -
accuracy: 0.9270 - loss: 1.0456 - val_accuracy: 0.9155 - val_loss: 1.1358
Epoch 73/100
```

```
895/895             0s 592ms/step -
accuracy: 0.9261 - loss: 1.0491
Epoch 73: val_loss did not improve from 1.13494
2025-10-23 18:21:11 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895             551s 615ms/step -
accuracy: 0.9261 - loss: 1.0491 - val_accuracy: 0.9152 - val_loss: 1.1378
Epoch 74/100
895/895             0s 591ms/step -
accuracy: 0.9248 - loss: 1.0545
Epoch 74: val_loss did not improve from 1.13494
2025-10-23 18:30:21 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895             550s 614ms/step -
accuracy: 0.9248 - loss: 1.0545 - val_accuracy: 0.9147 - val_loss: 1.1420
Epoch 75/100
895/895             0s 588ms/step -
accuracy: 0.9235 - loss: 1.0601
Epoch 75: val_loss did not improve from 1.13494
2025-10-23 18:39:29 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895             548s 612ms/step -
accuracy: 0.9235 - loss: 1.0601 - val_accuracy: 0.9141 - val_loss: 1.1459
Epoch 76/100
895/895             0s 589ms/step -
accuracy: 0.9225 - loss: 1.0650
Epoch 76: val_loss did not improve from 1.13494
2025-10-23 18:48:36 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895             548s 612ms/step -
accuracy: 0.9225 - loss: 1.0650 - val_accuracy: 0.9136 - val_loss: 1.1470
Epoch 76: early stopping
Restoring model weights from the end of the best epoch: 71.
2025-10-23 18:48:37 - bilstm_attention - INFO - Training completed in 0.93 hours
```

[14]:
```python
# Save model
model_path = f"{Config.ARTIFACT_PATH}/{model_name}"
os.makedirs(model_path, exist_ok=True)
trainer.save_model(f"{model_path}/final_bilstm_model.keras")
```

```
2025-10-23 18:48:38 - bilstm_attention - INFO - Model saved to
…/artifacts/bilstm/final_bilstm_model.keras
```

1.6 6. Build LSTM Model

[15]:
```
GPUMemoryManager.clear_session()

lstm_builder = LSTMAttentionModel(
    config=Config,
    name_logger="lstm_attention",
    filename_logger=f"{Config.LOG_DIR}/lstm_attention.log"
)
lstm_model = lstm_builder.build(
    vocab_size_src=Config.MAX_VOCAB_SIZE_SRC,
    vocab_size_trg=Config.MAX_VOCAB_SIZE_TRG,
    max_len_src=Config.MAX_LENGTH_SRC,
    max_len_trg=Config.MAX_LENGTH_TRG
)

lstm_model.summary()
```

```
Session cleared
```

Model: "lstm_attention"

Layer (type)	Output Shape	Param #	Connected to
encoder_input (InputLayer)	(None, 40)	0	-
decoder_input (InputLayer)	(None, 50)	0	-
encoder_embedding (Embedding)	(None, 40, 64)	1,920,000	encoder_input[0]…
decoder_embedding (Embedding)	(None, 50, 64)	1,600,000	decoder_input[0]…
encoder_ln1 (LayerNormalizatio…	(None, 40, 64)	128	encoder_embeddin…
decoder_ln1 (LayerNormalizatio…	(None, 50, 64)	128	decoder_embeddin…
encoder_dropout (Dropout)	(None, 40, 64)	0	encoder_ln1[0][0]
not_equal (NotEqual)	(None, 40)	0	encoder_input[0]…

decoder_dropout (Dropout)	(None, 50, 64)	0	decoder_ln1[0][0]
encoder_lstm (LSTM)	[(None, 40, 128), (None, 128), (None, 128)]	98,816	encoder_dropout[… not_equal[0][0]
decoder_lstm (LSTM)	[(None, 50, 128), (None, 128), (None, 128)]	98,816	decoder_dropout[… encoder_lstm[0][… encoder_lstm[0][…
decoder_ln2 (LayerNormalizatio…	(None, 50, 128)	256	decoder_lstm[0][…
encoder_ln2 (LayerNormalizatio…	(None, 40, 128)	256	encoder_lstm[0][…
not_equal_1 (NotEqual)	(None, 50)	0	decoder_input[0]…
attention_layer (MultiHeadAttentio…	(None, 50, 128)	263,808	encoder_ln2[0][0… not_equal[0][0], decoder_ln2[0][0… not_equal_1[0][0… encoder_ln2[0][0… not_equal[0][0]
attention_ln (LayerNormalizatio…	(None, 50, 128)	256	attention_layer[…
residual_add (Add)	(None, 50, 128)	0	decoder_ln2[0][0… attention_ln[0][…
residual_ln (LayerNormalizatio…	(None, 50, 128)	256	residual_add[0][…
concat_output (Concatenate)	(None, 50, 256)	0	residual_ln[0][0… attention_ln[0][…
concat_dropout (Dropout)	(None, 50, 256)	0	concat_output[0]…
decoder_dense (Dense)	(None, 50, 25000)	6,425,000	concat_dropout[0…

Total params: 10,407,720 (39.70 MB)

Trainable params: 10,407,720 (39.70 MB)

Non-trainable params: 0 (0.00 B)

1.7 7. Train LSTM

[16]:
```
model_name = "lstm"

lstm_trainer = ModelTrainer(
    model=lstm_model,
    config=Config,
    model_name=model_name,
    logger_name="lstm_attention",
    logger_file=f"{Config.LOG_DIR}/lstm_attention.log"
)

lstm_history = lstm_trainer.train(
    train_data=(en_train, vi_in_train, vi_out_train),
    val_data=(en_val, vi_in_val, vi_out_val),
    resume=True
)
```

```
2025-10-23 18:48:41 - lstm_attention - INFO - No checkpoint directory found.
Starting from scratch.
2025-10-23 18:48:41 - lstm_attention - INFO - Mixed precision optimizer enabled
with increased stability
2025-10-23 18:48:41 - lstm_attention - INFO - Model compiled:
2025-10-23 18:48:41 - lstm_attention - INFO - Loss: Label Smoothing ( =0.05)
2025-10-23 18:48:41 - lstm_attention - INFO - LR Schedule: True
2025-10-23 18:48:41 - lstm_attention - INFO - Training started:
2025-10-23 18:48:41 - lstm_attention - INFO - Batch size: 128
2025-10-23 18:48:41 - lstm_attention - INFO - Initial epoch: 0
2025-10-23 18:48:41 - lstm_attention - INFO - Total epochs: 100
Epoch 1/100
895/895            0s 388ms/step -
accuracy: 0.5073 - loss: 7.6140
Epoch 1: val_loss improved from inf to 2.50783, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 18:55:02 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            381s 409ms/step -
accuracy: 0.5075 - loss: 7.6113 - val_accuracy: 0.7143 - val_loss: 2.5078
Epoch 2/100
```

```
895/895            0s 387ms/step -
accuracy: 0.7280 - loss: 2.4225
Epoch 2: val_loss improved from 2.50783 to 2.03052, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 19:01:05 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            363s 405ms/step -
accuracy: 0.7280 - loss: 2.4224 - val_accuracy: 0.7826 - val_loss: 2.0305
Epoch 3/100
895/895            0s 388ms/step -
accuracy: 0.7837 - loss: 2.0286
Epoch 3: val_loss improved from 2.03052 to 1.71270, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 19:07:09 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            364s 407ms/step -
accuracy: 0.7837 - loss: 2.0285 - val_accuracy: 0.8266 - val_loss: 1.7127
Epoch 4/100
895/895            0s 388ms/step -
accuracy: 0.8179 - loss: 1.7698
Epoch 4: val_loss improved from 1.71270 to 1.52958, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 19:13:13 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            364s 407ms/step -
accuracy: 0.8179 - loss: 1.7697 - val_accuracy: 0.8513 - val_loss: 1.5296
Epoch 5/100
895/895            0s 389ms/step -
accuracy: 0.8385 - loss: 1.6117
Epoch 5: val_loss improved from 1.52958 to 1.44338, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 19:19:18 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            365s 408ms/step -
accuracy: 0.8385 - loss: 1.6117 - val_accuracy: 0.8635 - val_loss: 1.4434
Epoch 6/100
895/895            0s 386ms/step -
accuracy: 0.8509 - loss: 1.5160
Epoch 6: val_loss improved from 1.44338 to 1.37789, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 19:25:20 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            362s 405ms/step -
```

```
accuracy: 0.8509 - loss: 1.5160 - val_accuracy: 0.8721 - val_loss: 1.3779
Epoch 7/100
895/895            0s 387ms/step -
accuracy: 0.8591 - loss: 1.4556
Epoch 7: val_loss improved from 1.37789 to 1.34537, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 19:31:24 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            363s 406ms/step -
accuracy: 0.8591 - loss: 1.4556 - val_accuracy: 0.8772 - val_loss: 1.3454
Epoch 8/100
895/895            0s 387ms/step -
accuracy: 0.8652 - loss: 1.4119
Epoch 8: val_loss improved from 1.34537 to 1.31662, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 19:37:27 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            363s 406ms/step -
accuracy: 0.8652 - loss: 1.4119 - val_accuracy: 0.8816 - val_loss: 1.3166
Epoch 9/100
895/895            0s 389ms/step -
accuracy: 0.8697 - loss: 1.3792
Epoch 9: val_loss improved from 1.31662 to 1.30040, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 19:43:32 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            365s 408ms/step -
accuracy: 0.8697 - loss: 1.3792 - val_accuracy: 0.8838 - val_loss: 1.3004
Epoch 10/100
895/895            0s 389ms/step -
accuracy: 0.8730 - loss: 1.3565
Epoch 10: val_loss improved from 1.30040 to 1.28057, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 19:49:36 - src.training.callbacks - INFO - Saved checkpoint at epoch
10: /kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/checkpoint_epoch_010.keras
2025-10-23 19:49:38 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 409ms/step -
accuracy: 0.8730 - loss: 1.3565 - val_accuracy: 0.8869 - val_loss: 1.2806
Epoch 11/100
895/895            0s 389ms/step -
accuracy: 0.8772 - loss: 1.3290
Epoch 11: val_loss improved from 1.28057 to 1.26916, saving model to
```

```
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 19:55:43 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 408ms/step -
accuracy: 0.8772 - loss: 1.3290 - val_accuracy: 0.8889 - val_loss: 1.2692
Epoch 12/100
895/895            0s 388ms/step -
accuracy: 0.8795 - loss: 1.3133
Epoch 12: val_loss improved from 1.26916 to 1.25877, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 20:01:48 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            364s 407ms/step -
accuracy: 0.8795 - loss: 1.3133 - val_accuracy: 0.8908 - val_loss: 1.2588
Epoch 13/100
895/895            0s 390ms/step -
accuracy: 0.8821 - loss: 1.2971
Epoch 13: val_loss improved from 1.25877 to 1.25353, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 20:07:54 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 409ms/step -
accuracy: 0.8821 - loss: 1.2971 - val_accuracy: 0.8920 - val_loss: 1.2535
Epoch 14/100
895/895            0s 392ms/step -
accuracy: 0.8838 - loss: 1.2853
Epoch 14: val_loss improved from 1.25353 to 1.24509, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 20:14:01 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            367s 411ms/step -
accuracy: 0.8838 - loss: 1.2852 - val_accuracy: 0.8931 - val_loss: 1.2451
Epoch 15/100
895/895            0s 390ms/step -
accuracy: 0.8856 - loss: 1.2740
Epoch 15: val_loss improved from 1.24509 to 1.24039, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 20:20:07 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 409ms/step -
accuracy: 0.8856 - loss: 1.2740 - val_accuracy: 0.8940 - val_loss: 1.2404
Epoch 16/100
895/895            0s 392ms/step -
```

```
accuracy: 0.8867 - loss: 1.2656
Epoch 16: val_loss improved from 1.24039 to 1.23788, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 20:26:15 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            367s 411ms/step -
accuracy: 0.8867 - loss: 1.2656 - val_accuracy: 0.8949 - val_loss: 1.2379
Epoch 17/100
895/895            0s 391ms/step -
accuracy: 0.8882 - loss: 1.2571
Epoch 17: val_loss improved from 1.23788 to 1.23227, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 20:32:22 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            367s 410ms/step -
accuracy: 0.8882 - loss: 1.2571 - val_accuracy: 0.8957 - val_loss: 1.2323
Epoch 18/100
895/895            0s 391ms/step -
accuracy: 0.8893 - loss: 1.2514
Epoch 18: val_loss did not improve from 1.23227
2025-10-23 20:38:28 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 409ms/step -
accuracy: 0.8893 - loss: 1.2514 - val_accuracy: 0.8957 - val_loss: 1.2324
Epoch 19/100
895/895            0s 391ms/step -
accuracy: 0.8901 - loss: 1.2455
Epoch 19: val_loss improved from 1.23227 to 1.22595, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 20:44:35 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            367s 410ms/step -
accuracy: 0.8901 - loss: 1.2455 - val_accuracy: 0.8967 - val_loss: 1.2260
Epoch 20/100
895/895            0s 392ms/step -
accuracy: 0.8908 - loss: 1.2404
Epoch 20: val_loss did not improve from 1.22595
2025-10-23 20:50:41 - src.training.callbacks - INFO - Saved checkpoint at epoch
20: /kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/checkpoint_epoch_020.keras
2025-10-23 20:50:43 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            368s 411ms/step -
accuracy: 0.8908 - loss: 1.2404 - val_accuracy: 0.8968 - val_loss: 1.2269
Epoch 21/100
```

```
895/895            0s 391ms/step -
accuracy: 0.8917 - loss: 1.2360
Epoch 21: val_loss improved from 1.22595 to 1.22243, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 20:56:49 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            367s 410ms/step -
accuracy: 0.8917 - loss: 1.2360 - val_accuracy: 0.8973 - val_loss: 1.2224
Epoch 22/100
895/895            0s 391ms/step -
accuracy: 0.8922 - loss: 1.2322
Epoch 22: val_loss improved from 1.22243 to 1.22067, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 21:02:57 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            367s 410ms/step -
accuracy: 0.8922 - loss: 1.2322 - val_accuracy: 0.8981 - val_loss: 1.2207
Epoch 23/100
895/895            0s 390ms/step -
accuracy: 0.8932 - loss: 1.2267
Epoch 23: val_loss improved from 1.22067 to 1.21913, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 21:09:02 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 409ms/step -
accuracy: 0.8932 - loss: 1.2267 - val_accuracy: 0.8982 - val_loss: 1.2191
Epoch 24/100
895/895            0s 390ms/step -
accuracy: 0.8937 - loss: 1.2231
Epoch 24: val_loss improved from 1.21913 to 1.21700, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 21:15:08 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 409ms/step -
accuracy: 0.8937 - loss: 1.2231 - val_accuracy: 0.8991 - val_loss: 1.2170
Epoch 25/100
895/895            0s 390ms/step -
accuracy: 0.8942 - loss: 1.2207
Epoch 25: val_loss improved from 1.21700 to 1.21580, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 21:21:14 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 408ms/step -
```

```
accuracy: 0.8942 - loss: 1.2207 - val_accuracy: 0.8985 - val_loss: 1.2158
Epoch 26/100
895/895            0s 390ms/step -
accuracy: 0.8949 - loss: 1.2168
Epoch 26: val_loss did not improve from 1.21580
2025-10-23 21:27:19 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            365s 408ms/step -
accuracy: 0.8949 - loss: 1.2168 - val_accuracy: 0.8988 - val_loss: 1.2171
Epoch 27/100
895/895            0s 391ms/step -
accuracy: 0.8954 - loss: 1.2141
Epoch 27: val_loss improved from 1.21580 to 1.21275, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 21:33:26 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            367s 410ms/step -
accuracy: 0.8954 - loss: 1.2141 - val_accuracy: 0.8999 - val_loss: 1.2128
Epoch 28/100
895/895            0s 392ms/step -
accuracy: 0.8958 - loss: 1.2110
Epoch 28: val_loss improved from 1.21275 to 1.21149, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 21:39:34 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            368s 411ms/step -
accuracy: 0.8958 - loss: 1.2110 - val_accuracy: 0.8995 - val_loss: 1.2115
Epoch 29/100
895/895            0s 394ms/step -
accuracy: 0.8959 - loss: 1.2097
Epoch 29: val_loss improved from 1.21149 to 1.21107, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 21:45:43 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            370s 413ms/step -
accuracy: 0.8959 - loss: 1.2097 - val_accuracy: 0.9002 - val_loss: 1.2111
Epoch 30/100
895/895            0s 392ms/step -
accuracy: 0.8967 - loss: 1.2059
Epoch 30: val_loss improved from 1.21107 to 1.20955, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 21:51:50 - src.training.callbacks - INFO - Saved checkpoint at epoch
30: /kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/checkpoint_epoch_030.keras
```

```
2025-10-23 21:51:52 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            368s 412ms/step -
accuracy: 0.8967 - loss: 1.2059 - val_accuracy: 0.9004 - val_loss: 1.2095
Epoch 31/100
895/895            0s 389ms/step -
accuracy: 0.8970 - loss: 1.2030
Epoch 31: val_loss improved from 1.20955 to 1.20883, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 21:57:57 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 408ms/step -
accuracy: 0.8970 - loss: 1.2030 - val_accuracy: 0.9003 - val_loss: 1.2088
Epoch 32/100
895/895            0s 391ms/step -
accuracy: 0.8974 - loss: 1.2018
Epoch 32: val_loss improved from 1.20883 to 1.20605, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 22:04:05 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            367s 410ms/step -
accuracy: 0.8974 - loss: 1.2018 - val_accuracy: 0.9007 - val_loss: 1.2061
Epoch 33/100
895/895            0s 391ms/step -
accuracy: 0.8977 - loss: 1.1992
Epoch 33: val_loss improved from 1.20605 to 1.20528, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 22:10:12 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            368s 411ms/step -
accuracy: 0.8977 - loss: 1.1992 - val_accuracy: 0.9010 - val_loss: 1.2053
Epoch 34/100
895/895            0s 390ms/step -
accuracy: 0.8981 - loss: 1.1970
Epoch 34: val_loss did not improve from 1.20528
2025-10-23 22:16:18 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            365s 408ms/step -
accuracy: 0.8981 - loss: 1.1970 - val_accuracy: 0.9009 - val_loss: 1.2070
Epoch 35/100
895/895            0s 390ms/step -
accuracy: 0.8985 - loss: 1.1949
Epoch 35: val_loss did not improve from 1.20528
2025-10-23 22:22:23 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
```

```
895/895            366s 408ms/step -
accuracy: 0.8985 - loss: 1.1949 - val_accuracy: 0.9011 - val_loss: 1.2056
Epoch 36/100
895/895            0s 390ms/step -
accuracy: 0.8987 - loss: 1.1940
Epoch 36: val_loss improved from 1.20528 to 1.20394, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 22:28:30 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 409ms/step -
accuracy: 0.8987 - loss: 1.1940 - val_accuracy: 0.9015 - val_loss: 1.2039
Epoch 37/100
895/895            0s 390ms/step -
accuracy: 0.8988 - loss: 1.1925
Epoch 37: val_loss improved from 1.20394 to 1.20299, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 22:34:36 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 409ms/step -
accuracy: 0.8988 - loss: 1.1925 - val_accuracy: 0.9017 - val_loss: 1.2030
Epoch 38/100
895/895            0s 390ms/step -
accuracy: 0.8994 - loss: 1.1902
Epoch 38: val_loss did not improve from 1.20299
2025-10-23 22:40:41 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 408ms/step -
accuracy: 0.8994 - loss: 1.1902 - val_accuracy: 0.9014 - val_loss: 1.2047
Epoch 39/100
895/895            0s 390ms/step -
accuracy: 0.8996 - loss: 1.1886
Epoch 39: val_loss improved from 1.20299 to 1.20020, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 22:46:48 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 409ms/step -
accuracy: 0.8996 - loss: 1.1886 - val_accuracy: 0.9016 - val_loss: 1.2002
Epoch 40/100
895/895            0s 390ms/step -
accuracy: 0.8998 - loss: 1.1866
Epoch 40: val_loss did not improve from 1.20020
2025-10-23 22:52:52 - src.training.callbacks - INFO - Saved checkpoint at epoch
40: /kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/checkpoint_epoch_040.keras
2025-10-23 22:52:54 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
```

```
(93.2%) | Peak: 15269 MB
895/895            366s 409ms/step -
accuracy: 0.8998 - loss: 1.1866 - val_accuracy: 0.9016 - val_loss: 1.2023
Epoch 41/100
895/895            0s 507ms/step -
accuracy: 0.9005 - loss: 1.1829
Epoch 41: val_loss did not improve from 1.20020
2025-10-23 23:01:11 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            497s 556ms/step -
accuracy: 0.9005 - loss: 1.1829 - val_accuracy: 0.9022 - val_loss: 1.2004
Epoch 42/100
895/895            0s 980ms/step -
accuracy: 0.9007 - loss: 1.1823
Epoch 42: val_loss did not improve from 1.20020
2025-10-23 23:16:34 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            922s 1s/step -
accuracy: 0.9007 - loss: 1.1823 - val_accuracy: 0.9021 - val_loss: 1.2009
Epoch 43/100
895/895            0s 996ms/step -
accuracy: 0.9010 - loss: 1.1805
Epoch 43: val_loss did not improve from 1.20020
2025-10-23 23:32:13 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            939s 1s/step -
accuracy: 0.9010 - loss: 1.1805 - val_accuracy: 0.9020 - val_loss: 1.2003
Epoch 44/100
895/895            0s 1s/step -
accuracy: 0.9010 - loss: 1.1801
Epoch 44: val_loss improved from 1.20020 to 1.19936, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-23 23:48:14 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            961s 1s/step -
accuracy: 0.9010 - loss: 1.1801 - val_accuracy: 0.9024 - val_loss: 1.1994
Epoch 45/100
895/895            0s 1s/step -
accuracy: 0.9013 - loss: 1.1788
Epoch 45: val_loss improved from 1.19936 to 1.19714, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 00:04:16 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            963s 1s/step -
accuracy: 0.9013 - loss: 1.1788 - val_accuracy: 0.9029 - val_loss: 1.1971
Epoch 46/100
```

```
895/895            0s 995ms/step -
accuracy: 0.9015 - loss: 1.1772
Epoch 46: val_loss did not improve from 1.19714
2025-10-24 00:19:54 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            938s 1s/step -
accuracy: 0.9015 - loss: 1.1772 - val_accuracy: 0.9025 - val_loss: 1.1975
Epoch 47/100
895/895            0s 1s/step -
accuracy: 0.9018 - loss: 1.1754
Epoch 47: val_loss improved from 1.19714 to 1.19523, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 00:35:58 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            964s 1s/step -
accuracy: 0.9018 - loss: 1.1754 - val_accuracy: 0.9030 - val_loss: 1.1952
Epoch 48/100
895/895            0s 975ms/step -
accuracy: 0.9021 - loss: 1.1741
Epoch 48: val_loss did not improve from 1.19523
2025-10-24 00:51:14 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            916s 1s/step -
accuracy: 0.9021 - loss: 1.1741 - val_accuracy: 0.9029 - val_loss: 1.1958
Epoch 49/100
895/895            0s 883ms/step -
accuracy: 0.9023 - loss: 1.1725
Epoch 49: val_loss improved from 1.19523 to 1.19503, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 01:04:56 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            822s 918ms/step -
accuracy: 0.9023 - loss: 1.1725 - val_accuracy: 0.9030 - val_loss: 1.1950
Epoch 50/100
895/895            0s 505ms/step -
accuracy: 0.9027 - loss: 1.1717
Epoch 50: val_loss did not improve from 1.19503
2025-10-24 01:12:44 - src.training.callbacks - INFO - Saved checkpoint at epoch
50: /kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/checkpoint_epoch_050.keras
2025-10-24 01:12:45 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            469s 524ms/step -
accuracy: 0.9027 - loss: 1.1717 - val_accuracy: 0.9029 - val_loss: 1.1952
Epoch 51/100
895/895            0s 395ms/step -
```

```
accuracy: 0.9029 - loss: 1.1697
Epoch 51: val_loss improved from 1.19503 to 1.19326, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 01:18:56 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            371s 414ms/step -
accuracy: 0.9029 - loss: 1.1697 - val_accuracy: 0.9033 - val_loss: 1.1933
Epoch 52/100
895/895            0s 394ms/step -
accuracy: 0.9030 - loss: 1.1689
Epoch 52: val_loss did not improve from 1.19326
2025-10-24 01:25:05 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            369s 412ms/step -
accuracy: 0.9030 - loss: 1.1689 - val_accuracy: 0.9036 - val_loss: 1.1944
Epoch 53/100
895/895            0s 394ms/step -
accuracy: 0.9034 - loss: 1.1676
Epoch 53: val_loss improved from 1.19326 to 1.19323, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 01:31:15 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            370s 413ms/step -
accuracy: 0.9034 - loss: 1.1676 - val_accuracy: 0.9036 - val_loss: 1.1932
Epoch 54/100
895/895            0s 395ms/step -
accuracy: 0.9034 - loss: 1.1673
Epoch 54: val_loss did not improve from 1.19323
2025-10-24 01:37:25 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            370s 413ms/step -
accuracy: 0.9034 - loss: 1.1673 - val_accuracy: 0.9034 - val_loss: 1.1933
Epoch 55/100
895/895            0s 403ms/step -
accuracy: 0.9039 - loss: 1.1651
Epoch 55: val_loss improved from 1.19323 to 1.19232, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 01:43:43 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            378s 423ms/step -
accuracy: 0.9039 - loss: 1.1651 - val_accuracy: 0.9038 - val_loss: 1.1923
Epoch 56/100
895/895            0s 396ms/step -
accuracy: 0.9039 - loss: 1.1642
Epoch 56: val_loss did not improve from 1.19232
```

```
2025-10-24 01:49:54 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            371s 414ms/step -
accuracy: 0.9039 - loss: 1.1642 - val_accuracy: 0.9037 - val_loss: 1.1926
Epoch 57/100
895/895            0s 394ms/step -
accuracy: 0.9039 - loss: 1.1645
Epoch 57: val_loss did not improve from 1.19232
2025-10-24 01:56:03 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            369s 412ms/step -
accuracy: 0.9039 - loss: 1.1645 - val_accuracy: 0.9038 - val_loss: 1.1934
Epoch 58/100
895/895            0s 394ms/step -
accuracy: 0.9044 - loss: 1.1613
Epoch 58: val_loss improved from 1.19232 to 1.19144, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 02:02:13 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            370s 413ms/step -
accuracy: 0.9044 - loss: 1.1613 - val_accuracy: 0.9038 - val_loss: 1.1914
Epoch 59/100
895/895            0s 394ms/step -
accuracy: 0.9044 - loss: 1.1615
Epoch 59: val_loss did not improve from 1.19144
2025-10-24 02:08:22 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            369s 412ms/step -
accuracy: 0.9044 - loss: 1.1615 - val_accuracy: 0.9038 - val_loss: 1.1918
Epoch 60/100
895/895            0s 394ms/step -
accuracy: 0.9049 - loss: 1.1595
Epoch 60: val_loss did not improve from 1.19144
2025-10-24 02:14:30 - src.training.callbacks - INFO - Saved checkpoint at epoch
60: /kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/checkpoint_epoch_060.keras
2025-10-24 02:14:32 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            370s 413ms/step -
accuracy: 0.9049 - loss: 1.1595 - val_accuracy: 0.9039 - val_loss: 1.1916
Epoch 61/100
895/895            0s 396ms/step -
accuracy: 0.9050 - loss: 1.1585
Epoch 61: val_loss improved from 1.19144 to 1.18986, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 02:20:44 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
```

```
(93.2%) | Peak: 15269 MB
895/895            372s 416ms/step -
accuracy: 0.9050 - loss: 1.1585 - val_accuracy: 0.9043 - val_loss: 1.1899
Epoch 62/100
895/895            0s 415ms/step -
accuracy: 0.9051 - loss: 1.1581
Epoch 62: val_loss did not improve from 1.18986
2025-10-24 02:27:11 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            387s 432ms/step -
accuracy: 0.9051 - loss: 1.1581 - val_accuracy: 0.9042 - val_loss: 1.1902
Epoch 63/100
895/895            0s 395ms/step -
accuracy: 0.9053 - loss: 1.1571
Epoch 63: val_loss improved from 1.18986 to 1.18859, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 02:33:21 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            370s 414ms/step -
accuracy: 0.9053 - loss: 1.1571 - val_accuracy: 0.9045 - val_loss: 1.1886
Epoch 64/100
895/895            0s 392ms/step -
accuracy: 0.9056 - loss: 1.1554
Epoch 64: val_loss did not improve from 1.18859
2025-10-24 02:39:28 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            367s 410ms/step -
accuracy: 0.9056 - loss: 1.1554 - val_accuracy: 0.9045 - val_loss: 1.1896
Epoch 65/100
895/895            0s 391ms/step -
accuracy: 0.9059 - loss: 1.1540
Epoch 65: val_loss did not improve from 1.18859
2025-10-24 02:45:34 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 409ms/step -
accuracy: 0.9059 - loss: 1.1540 - val_accuracy: 0.9044 - val_loss: 1.1889
Epoch 66/100
895/895            0s 390ms/step -
accuracy: 0.9059 - loss: 1.1542
Epoch 66: val_loss improved from 1.18859 to 1.18796, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 02:51:40 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            366s 409ms/step -
accuracy: 0.9059 - loss: 1.1542 - val_accuracy: 0.9045 - val_loss: 1.1880
Epoch 67/100
```

```
895/895            0s 392ms/step -
accuracy: 0.9062 - loss: 1.1525
Epoch 67: val_loss improved from 1.18796 to 1.18712, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 02:57:48 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            368s 411ms/step -
accuracy: 0.9062 - loss: 1.1525 - val_accuracy: 0.9046 - val_loss: 1.1871
Epoch 68/100
895/895            0s 392ms/step -
accuracy: 0.9064 - loss: 1.1519
Epoch 68: val_loss did not improve from 1.18712
2025-10-24 03:03:55 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            367s 410ms/step -
accuracy: 0.9064 - loss: 1.1519 - val_accuracy: 0.9048 - val_loss: 1.1872
Epoch 69/100
895/895            0s 392ms/step -
accuracy: 0.9065 - loss: 1.1512
Epoch 69: val_loss improved from 1.18712 to 1.18622, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 03:10:02 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            367s 410ms/step -
accuracy: 0.9065 - loss: 1.1512 - val_accuracy: 0.9048 - val_loss: 1.1862
Epoch 70/100
895/895            0s 392ms/step -
accuracy: 0.9066 - loss: 1.1505
Epoch 70: val_loss did not improve from 1.18622
2025-10-24 03:16:08 - src.training.callbacks - INFO - Saved checkpoint at epoch
70: /kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/checkpoint_epoch_070.keras
2025-10-24 03:16:10 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            368s 411ms/step -
accuracy: 0.9066 - loss: 1.1505 - val_accuracy: 0.9048 - val_loss: 1.1869
Epoch 71/100
895/895            0s 392ms/step -
accuracy: 0.9066 - loss: 1.1496
Epoch 71: val_loss did not improve from 1.18622
2025-10-24 03:22:18 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            367s 411ms/step -
accuracy: 0.9066 - loss: 1.1496 - val_accuracy: 0.9049 - val_loss: 1.1865
Epoch 72/100
895/895            0s 392ms/step -
```

```
accuracy: 0.9071 - loss: 1.1476
Epoch 72: val_loss improved from 1.18622 to 1.18583, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 03:28:25 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            368s 411ms/step -
accuracy: 0.9071 - loss: 1.1476 - val_accuracy: 0.9050 - val_loss: 1.1858
Epoch 73/100
895/895            0s 397ms/step -
accuracy: 0.9073 - loss: 1.1475
Epoch 73: val_loss did not improve from 1.18583
2025-10-24 03:34:37 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            371s 415ms/step -
accuracy: 0.9073 - loss: 1.1475 - val_accuracy: 0.9049 - val_loss: 1.1865
Epoch 74/100
895/895            0s 393ms/step -
accuracy: 0.9073 - loss: 1.1466
Epoch 74: val_loss did not improve from 1.18583
2025-10-24 03:40:45 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            368s 412ms/step -
accuracy: 0.9073 - loss: 1.1466 - val_accuracy: 0.9052 - val_loss: 1.1863
Epoch 75/100
895/895            0s 398ms/step -
accuracy: 0.9074 - loss: 1.1464
Epoch 75: val_loss did not improve from 1.18583
2025-10-24 03:46:57 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            372s 416ms/step -
accuracy: 0.9074 - loss: 1.1464 - val_accuracy: 0.9052 - val_loss: 1.1859
Epoch 76/100
895/895            0s 399ms/step -
accuracy: 0.9078 - loss: 1.1446
Epoch 76: val_loss improved from 1.18583 to 1.18539, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 03:53:11 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            374s 418ms/step -
accuracy: 0.9078 - loss: 1.1446 - val_accuracy: 0.9054 - val_loss: 1.1854
Epoch 77/100
895/895            0s 395ms/step -
accuracy: 0.9080 - loss: 1.1446
Epoch 77: val_loss did not improve from 1.18539
2025-10-24 03:59:21 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
```

```
895/895            370s 413ms/step -
accuracy: 0.9080 - loss: 1.1446 - val_accuracy: 0.9053 - val_loss: 1.1863
Epoch 78/100
895/895            0s 398ms/step -
accuracy: 0.9083 - loss: 1.1431
Epoch 78: val_loss improved from 1.18539 to 1.18528, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 04:05:34 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            373s 417ms/step -
accuracy: 0.9083 - loss: 1.1431 - val_accuracy: 0.9053 - val_loss: 1.1853
Epoch 79/100
895/895            0s 396ms/step -
accuracy: 0.9082 - loss: 1.1431
Epoch 79: val_loss improved from 1.18528 to 1.18457, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 04:11:46 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            372s 415ms/step -
accuracy: 0.9082 - loss: 1.1431 - val_accuracy: 0.9054 - val_loss: 1.1846
Epoch 80/100
895/895            0s 398ms/step -
accuracy: 0.9086 - loss: 1.1414
Epoch 80: val_loss improved from 1.18457 to 1.18443, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 04:17:58 - src.training.callbacks - INFO - Saved checkpoint at epoch
80: /kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/checkpoint_epoch_080.keras
2025-10-24 04:18:00 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            374s 418ms/step -
accuracy: 0.9086 - loss: 1.1414 - val_accuracy: 0.9054 - val_loss: 1.1844
Epoch 81/100
895/895            0s 397ms/step -
accuracy: 0.9085 - loss: 1.1408
Epoch 81: val_loss improved from 1.18443 to 1.18430, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 04:24:12 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            372s 416ms/step -
accuracy: 0.9085 - loss: 1.1408 - val_accuracy: 0.9057 - val_loss: 1.1843
Epoch 82/100
895/895            0s 397ms/step -
accuracy: 0.9086 - loss: 1.1415
```

```
Epoch 82: val_loss improved from 1.18430 to 1.18401, saving model to
/kaggle/working/machine-learning-
resource/artifacts/lstm/checkpoints/best_model.keras
2025-10-24 04:30:25 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            372s 416ms/step -
accuracy: 0.9086 - loss: 1.1415 - val_accuracy: 0.9056 - val_loss: 1.1840
Epoch 83/100
895/895            0s 397ms/step -
accuracy: 0.9089 - loss: 1.1402
Epoch 83: val_loss did not improve from 1.18401
2025-10-24 04:36:37 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            372s 416ms/step -
accuracy: 0.9089 - loss: 1.1402 - val_accuracy: 0.9054 - val_loss: 1.1852
Epoch 84/100
895/895            0s 401ms/step -
accuracy: 0.9090 - loss: 1.1398
Epoch 84: val_loss did not improve from 1.18401
2025-10-24 04:42:51 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            375s 419ms/step -
accuracy: 0.9090 - loss: 1.1398 - val_accuracy: 0.9056 - val_loss: 1.1849
Epoch 85/100
895/895            0s 400ms/step -
accuracy: 0.9091 - loss: 1.1393
Epoch 85: val_loss did not improve from 1.18401
2025-10-24 04:49:05 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            374s 418ms/step -
accuracy: 0.9091 - loss: 1.1393 - val_accuracy: 0.9057 - val_loss: 1.1852
Epoch 86/100
895/895            0s 401ms/step -
accuracy: 0.9091 - loss: 1.1392
Epoch 86: val_loss did not improve from 1.18401
2025-10-24 04:55:21 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            375s 419ms/step -
accuracy: 0.9091 - loss: 1.1392 - val_accuracy: 0.9056 - val_loss: 1.1855
Epoch 87/100
895/895            0s 398ms/step -
accuracy: 0.9093 - loss: 1.1385
Epoch 87: val_loss did not improve from 1.18401
2025-10-24 05:01:33 - src.training.callbacks - INFO - GPU Memory: 15269/16384 MB
(93.2%) | Peak: 15269 MB
895/895            373s 416ms/step -
accuracy: 0.9093 - loss: 1.1385 - val_accuracy: 0.9056 - val_loss: 1.1859
Epoch 87: early stopping
```

```
Restoring model weights from the end of the best epoch: 82.
2025-10-24 05:01:33 - lstm_attention - INFO - Training completed in 10.21 hours
```

[17]:
```python
# Save
model_path = f"{Config.ARTIFACT_PATH}/{model_name}"
os.makedirs(model_path, exist_ok=True)
lstm_trainer.save_model(f"{model_path}/final_lstm_model.keras")
```

```
2025-10-24 05:01:42 - lstm_attention - INFO - Model saved to
…/artifacts/lstm/final_lstm_model.keras
```

1.8 8. Compare Results

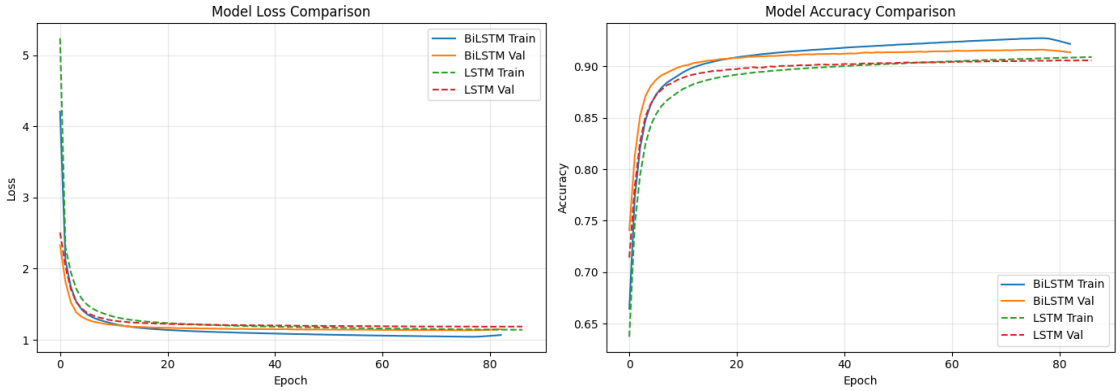
[18]:
```python
fig, axes = plt.subplots(1, 2, figsize=(14, 5))

# Loss
axes[0].plot(bilstm_history.history['loss'], label='BiLSTM Train')
axes[0].plot(bilstm_history.history['val_loss'], label='BiLSTM Val')
axes[0].plot(lstm_history.history['loss'], label='LSTM Train', linestyle='--')
axes[0].plot(lstm_history.history['val_loss'], label='LSTM Val', linestyle='--')
axes[0].set_title('Model Loss Comparison')
axes[0].set_xlabel('Epoch')
axes[0].set_ylabel('Loss')
axes[0].legend()
axes[0].grid(True, alpha=0.3)

# Accuracy
axes[1].plot(bilstm_history.history['accuracy'], label='BiLSTM Train')
axes[1].plot(bilstm_history.history['val_accuracy'], label='BiLSTM Val')
axes[1].plot(lstm_history.history['accuracy'], label='LSTM Train',␣
 ↪linestyle='--')
axes[1].plot(lstm_history.history['val_accuracy'], label='LSTM Val',␣
 ↪linestyle='--')
axes[1].set_title('Model Accuracy Comparison')
axes[1].set_xlabel('Epoch')
axes[1].set_ylabel('Accuracy')
axes[1].legend()
axes[1].grid(True, alpha=0.3)

plt.tight_layout()

# Save figure
os.makedirs(Config.ASSETS_PATH, exist_ok=True)
plt.savefig(f"{Config.ASSETS_PATH}/comparison.png", dpi=300)
plt.show()
```



[19]:
```
print(f"BiLSTM - Final Val Loss: {bilstm_history.history['val_loss'][-1]:.4f}")
print(f"LSTM - Final Val Loss: {lstm_history.history['val_loss'][-1]:.4f}")
```

```
BiLSTM - Final Val Loss: 1.1470
LSTM - Final Val Loss: 1.1859
```